



OIPE

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/045,072

DATE: 01/28/2002

TIME: 14:08:17

Input Set : A:\seqlist-15330790002 ascii

Output Set: N:\CRF3\01282002\J045072.raw

7 <110> APPLICANT: Sinskey, Anthony J.
 10 Lessard, Philip A.
 13 Willis, Laura B.
 22 <120> TITLE OF INVENTION: Pyruvate Carboxylase from Corynebacterium glutamicum
 28 <130> FILE REFERENCE: 1533.0790002
 33 <140> CURRENT APPLICATION NUMBER: US/10/045,072
 33 <141> CURRENT FILING DATE: 2002-01-15
 33 <150> PRIOR APPLICATION NUMBER: US 09/677,575
 36 <151> PRIOR FILING DATE: 2000-10-03
 42 <150> PRIOR APPLICATION NUMBER: US 09/220,081
 45 <151> PRIOR FILING DATE: 1998-12-23
 51 <160> NUMBER OF SEQ ID NOS: 39
 57 <170> SOFTWARE: PatentIn Ver. 2.0
 63 <210> SEQ ID NO: 1
 66 <211> LENGTH: 3621
 69 <212> TYPE: DNA
 72 <213> ORGANISM: Corynebacterium glutamicum
 78 <220> FEATURE:
 81 <221> NAME/KEY: CDS
 84 <222> LOCATION: (199)..(3621)
 90 <400> SEQUENCE: 1
 93 tggggcgggg ttagatcctg ggggggtttat ttcattcact ttggcttgaa gtcgtgcagg 60
 99 tcaggggaggt gttgcccgaa aacattgaga ggaaacaaa aaccgatgtt tgattggggg 120
 105 aatcggggggt tacgatacta ggacgcagtg actgctatca cccttggcgg tctcttggtg 180
 111 aaaggaataa ttactcta gtg tcg act cac aca tct tca acg ctt cca gca 231
 114 Met Ser Thr His Thr Ser Ser Thr Leu Pro Ala
 117 1 5 10
 123 ttc aaa aag atc ttg gta gca aac cgc ggc gaa atc gcg gtc cgt gct 279
 126 Phe Lys Lys Ile Leu Val Ala Asn Arg Gly Glu Ile Ala Val Arg Ala
 129 15 20 25
 135 ttc cgt gca gca ctc gaa acc ggt gca gcc acg gta gct att tac ccc 327
 138 Phe Arg Ala Ala Leu Glu Thr Gly Ala Ala Thr Val Ala Ile Tyr Pro
 141 30 35 40
 147 cgt gaa gat cgg gga tca ttc cac cgc tct ttt gct tct gaa gct gtc 375
 150 Arg Glu Asp Arg Gly Ser Phe His Arg Ser Phe Ala Ser Glu Ala Val
 153 45 50 55
 159 cgc att ggt acc gaa ggc tca cca gtc aag gcg tac ctg gac atc gat 423
 162 Arg Ile Gly Thr Glu Gly Ser Pro Val Lys Ala Tyr Leu Asp Ile Asp
 165 60 65 70 75
 171 gaa att atc ggt gca gct aaa aaa gtt aaa gca gat gcc att tac ccg 471
 174 Glu Ile Ile Gly Ala Ala Lys Lys Val Lys Ala Asp Ala Ile Tyr Pro
 177 80 85 90
 183 gga tac ggc ttc ctg tct gaa aat gcc cag ctt gcc cgc gag tgt gcg 519

ENTERED

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/045,072

DATE: 01/28/2002

TIME: 14:08:17

Input Set : A:\seqlist-15330790002 ascii

Output Set: N:\CRF3\01282002\J045072.raw

```

186 Gly Tyr Gly Phe Leu Ser Glu Asn Ala Gln Leu Ala Arg Glu Cys Ala
189          95          100          105
195 gaa aac ggc att act ttt att ggc cca acc cca gag gtt ctt gat ctc 567
198 Glu Asn Gly Ile Thr Phe Ile Gly Pro Thr Pro Glu Val Leu Asp Leu
201          110          115          120
207 acc ggt gat aag tct cgc gcg gta acc gcc gcg aag aag gct ggt ctg 615
210 Thr Gly Asp Lys Ser Arg Ala Val Thr Ala Ala Lys Lys Ala Gly Leu
213          125          130          135
219 cca gtt ttg gcg gaa tcc acc ccg agc aaa aac atc gat gag atc gtt 663
222 Pro Val Leu Ala Glu Ser Thr Pro Ser Lys Asn Ile Asp Glu Ile Val
225 140          145          150          155
231 aaa agc gct gaa ggc cag act tac ccc atc ttt gtg aag gca gtt gcc 711
234 Lys Ser Ala Glu Gly Gln Thr Tyr Pro Ile Phe Val Lys Ala Val Ala
237          160          165          170
243 ggt ggt ggc gga cgc ggt atg cgt ttt gtt gct tca cct gat gag ctt 759
246 Gly Gly Gly Gly Arg Gly Met Arg Phe Val Ala Ser Pro Asp Glu Leu
249          175          180          185
255 cgc aaa tta gca aca gaa gca tct cgt gaa gct gaa gcg gct ttc ggc 807
258 Arg Lys Leu Ala Thr Glu Ala Ser Arg Glu Ala Glu Ala Ala Phe Gly
261          190          195          200
267 gat ggc gcg gta tat gtc gaa cgt gct gtg att aac cct cag cat att 855
270 Asp Gly Ala Val Tyr Val Glu Arg Ala Val Ile Asn Pro Gln His Ile
273          205          210          215
279 gaa gtg cag atc ctt ggc gat cac act gga gaa gtt gta cac ctt tat 903
282 Glu Val Gln Ile Leu Gly Asp His Thr Gly Glu Val Val His Leu Tyr
285 220          225          230          235
291 gaa cgt gac tgc tca ctg cag cgt cgt cac caa aaa gtt gtc gaa att 951
294 Glu Arg Asp Cys Ser Leu Gln Arg Arg His Gln Lys Val Val Glu Ile
297          240          245          250
303 gcg cca gca cag cat ttg gat cca gaa ctg cgt gat cgc att tgt gcg 999
306 Ala Pro Ala Gln His Leu Asp Pro Glu Leu Arg Asp Arg Ile Cys Ala
309          255          260          265
315 gat gca gta aag ttc tgc cgc tcc att ggt tac cag ggc gcg gga acc 1047
318 Asp Ala Val Lys Phe Cys Arg Ser Ile Gly Tyr Gln Gly Ala Gly Thr
321          270          275          280
327 gtg gaa ttc ttg gtc gat gaa aag ggc aac cac gtc ttc atc gaa atg 1095
330 Val Glu Phe Leu Val Asp Glu Lys Gly Asn His Val Phe Ile Glu Met
333          285          290          295
339 aac cca cgt atc cag gtt gag cac acc gtg act gaa gaa gtc acc gag 1143
342 Asn Pro Arg Ile Gln Val Glu His Thr Val Thr Glu Glu Val Thr Glu
345 300          305          310          315
351 gtg gac ctg gtg aag gcg cag atg cgc ttg gct gct ggt gca acc ttg 1191
354 Val Asp Leu Val Lys Ala Gln Met Arg Leu Ala Ala Gly Ala Thr Leu
357          320          325          330
363 aag gaa ttg ggt ctg acc caa gat aag atc aag acc cac ggt gca gca 1239
366 Lys Glu Leu Gly Leu Thr Gln Asp Lys Ile Lys Thr His Gly Ala Ala
369          335          340          345
375 ctg cag tgc cgc atc acc acg gaa gat cca aac aac ggc ttc cgc cca 1287
378 Leu Gln Cys Arg Ile Thr Thr Glu Asp Pro Asn Asn Gly Phe Arg Pro

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/045,072

DATE: 01/28/2002

TIME: 14:08:17

Input Set : A:\seqlist-15330790002 ascii

Output Set: N:\CRF3\01282002\J045072.raw

```

381          350          355          360
387 gat acc gga act atc acc gcg tac cgc tca cca ggc gga gct ggc gtt 1335
390 Asp Thr Gly Thr Ile Thr Ala Tyr Arg Ser Pro Gly Gly Ala Gly Val
393          365          370          375
399 cgt ctt gac ggt gca gct cag ctc ggt ggc gaa atc acc gca cac ttt 1383
402 Arg Leu Asp Gly Ala Ala Gln Leu Gly Gly Glu Ile Thr Ala His Phe
405 380          385          390          395
411 gac tcc atg ctg gtg aaa atg acc tgc cgt ggt tcc gac ttt gaa act 1431
414 Asp Ser Met Leu Val Lys Met Thr Cys Arg Gly Ser Asp Phe Glu Thr
417          400          405          410
423 gct gtt gct cgt gca cag cgc gcg ttg gct gag ttc acc gtg tct ggt 1479
426 Ala Val Ala Arg Ala Gln Arg Ala Leu Ala Glu Phe Thr Val Ser Gly
429          415          420          425
435 gtt gca acc aac att ggt ttc ttg cgt gcg ttg ctg cgg gaa gag gac 1527
438 Val Ala Thr Asn Ile Gly Phe Leu Arg Ala Leu Leu Arg Glu Glu Asp
441          430          435          440
447 ttc act tcc aag cgc atc gcc acc gga ttc att gcc gat cac ccg cac 1575
450 Phe Thr Ser Lys Arg Ile Ala Thr Gly Phe Ile Ala Asp His Pro His
453          445          450          455
459 ctc ctt cag gct cca cct gct gat gat gag cag gga cgc atc ctg gat 1623
462 Leu Leu Gln Ala Pro Pro Ala Asp Asp Glu Gln Gly Arg Ile Leu Asp
465 460          465          470          475
471 tac ttg gca gat gtc acc gtg aac aag cct cat ggt gtg cgt cca aag 1671
474 Tyr Leu Ala Asp Val Thr Val Asn Lys Pro His Gly Val Arg Pro Lys
477          480          485          490
483 gat gtt gca gct cct atc gat aag ctg cct aac atc aag gat ctg cca 1719
486 Asp Val Ala Ala Pro Ile Asp Lys Leu Pro Asn Ile Lys Asp Leu Pro
489          495          500          505
495 ctg cca cgc ggt tcc cgt gac cgc ctg aag cag ctt ggc cca gcc gcg 1767
498 Leu Pro Arg Gly Ser Arg Asp Arg Leu Lys Gln Leu Gly Pro Ala Ala
501          510          515          520
507 ttt gct cgt gat ctc cgt gag cag gac gca ctg gca gtt act gat acc 1815
510 Phe Ala Arg Asp Leu Arg Glu Gln Asp Ala Leu Ala Val Thr Asp Thr
513          525          530          535
519 acc ttc cgc gat gca cac cag tct ttg ctt gcg acc cga gtc cgc tca 1863
522 Thr Phe Arg Asp Ala His Gln Ser Leu Leu Ala Thr Arg Val Arg Ser
525 540          545          550          555
531 ttc gca ctg aag cct gcg gca gag gcc gtc gca aag ctg act cct gag 1911
534 Phe Ala Leu Lys Pro Ala Ala Glu Ala Val Ala Lys Leu Thr Pro Glu
537          560          565          570
543 ctt ttg tcc gtg gag gcc tgg ggc ggc gcg acc tac gat gtg gcg atg 1959
546 Leu Leu Ser Val Glu Ala Trp Gly Gly Ala Thr Tyr Asp Val Ala Met
549          575          580          585
555 cgt ttc ctc ttt gag gat ccg tgg gac agg ctc gac gag ctg cgc gag 2007
558 Arg Phe Leu Phe Glu Asp Pro Trp Asp Arg Leu Asp Glu Leu Arg Glu
561          590          595          600
567 gcg atg ccg aat gta aac att cag atg ctg ctt cgc ggc cgc aac acc 2055
570 Ala Met Pro Asn Val Asn Ile Gln Met Leu Leu Arg Gly Arg Asn Thr
573          605          610          615

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/045,072

DATE: 01/28/2002

TIME: 14:08:17

Input Set : A:\seqlist-15330790002 ascii

Output Set: N:\CRF3\01282002\J045072.raw

```

579 gtg gga tac acc ccg tac cca gac tcc gtc tgc cgc gcg ttt gtt aag 2103
582 Val Gly Tyr Thr Pro Tyr Pro Asp Ser Val Cys Arg Ala Phe Val Lys
585 620 625 630 635
591 gaa gct gcc agc tcc ggc gtg gac atc ttc cgc atc ttc gac gcg ctt 2151
594 Glu Ala Ala Ser Ser Gly Val Asp Ile Phe Arg Ile Phe Asp Ala Leu
597 640 645 650
603 aac gac gtc tcc cag atg cgt cca gca atc gac gca gtc ctg gag acc 2199
606 Asn Asp Val Ser Gln Met Arg Pro Ala Ile Asp Ala Val Leu Glu Thr
609 655 660 665
615 aac acc gcg gta gcc gag gtg gct atg gct tat tct ggt gat ctc tct 2247
618 Asn Thr Ala Val Ala Glu Val Ala Met Ala Tyr Ser Gly Asp Leu Ser
621 670 675 680
627 gat cca aat gaa aag ctc tac acc ctg gat tac tac cta aag atg gca 2295
630 Asp Pro Asn Glu Lys Leu Tyr Thr Leu Asp Tyr Tyr Leu Lys Met Ala
633 685 690 695
639 gag gag atc gtc aag tct ggc gct cac atc ttg gcc att aag gat atg 2343
642 Glu Glu Ile Val Lys Ser Gly Ala His Ile Leu Ala Ile Lys Asp Met
645 700 705 710 715
651 gct ggt ctg ctt cgc cca gct gcg gta acc aag ctg gtc acc gca ctg 2391
654 Ala Gly Leu Leu Arg Pro Ala Ala Val Thr Lys Leu Val Thr Ala Leu
657 720 725 730
663 cgc cgt gaa ttc gat ctg cca gtg cac gtg cac acc cac gac act gcg 2439
666 Arg Arg Glu Phe Asp Leu Pro Val His Val His Thr His Asp Thr Ala
669 735 740 745
675 ggt ggc cag ctg gca acc tac ttt gct gca gct caa gct ggt gca gat 2487
678 Gly Gly Gln Leu Ala Thr Tyr Phe Ala Ala Ala Gln Ala Gly Ala Asp
681 750 755 760
687 gct gtt gac ggt gct tcc gca cca ctg tct ggc acc acc tcc cag cca 2535
690 Ala Val Asp Gly Ala Ser Ala Pro Leu Ser Gly Thr Thr Ser Gln Pro
693 765 770 775
699 tcc ctg tct gcc att gtt gct gca ttc gcg cac acc cgt cgc gat acc 2583
702 Ser Leu Ser Ala Ile Val Ala Ala Phe Ala His Thr Arg Arg Asp Thr
705 780 785 790 795
711 ggt ttg agc ctc gag gct gtt tct gac ctc gag ccg tac tgg gaa gca 2631
714 Gly Leu Ser Leu Glu Ala Val Ser Asp Leu Glu Pro Tyr Trp Glu Ala
717 800 805 810
723 gtg cgc gga ctg tac ctg cca ttt gag tct gga acc cca ggc cca acc 2679
726 Val Arg Gly Leu Tyr Leu Pro Phe Glu Ser Gly Thr Pro Gly Pro Thr
729 815 820 825
735 ggt cgc gtc tac cgc cac gaa atc cca ggc gga cag ttg tcc aac ctg 2727
738 Gly Arg Val Tyr Arg His Glu Ile Pro Gly Gly Gln Leu Ser Asn Leu
741 830 835 840
747 cgt gca cag gcc acc gca ctg ggc ctt gcg gat cgt ttc gaa ctc atc 2775
750 Arg Ala Gln Ala Thr Ala Leu Gly Leu Ala Asp Arg Phe Glu Leu Ile
753 845 850 855
759 gaa gac aac tac gca gcc gtt aat gag atg ctg gga cgc cca acc aag 2823
762 Glu Asp Asn Tyr Ala Ala Val Asn Glu Met Leu Gly Arg Pro Thr Lys
765 860 865 870 875
771 gtc acc cca tcc tcc aag gtt gtt ggc gac ctc gca ctc cac ctc gtt 2871

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/045,072

DATE: 01/28/2002

TIME: 14:08:17

Input Set : A:\seqlist-15330790002 ascii

Output Set: N:\CRF3\01282002\J045072.raw

```

774 Val Thr Pro Ser Ser Lys Val Val Gly Asp Leu Ala Leu His Leu Val
777      880      885      890
783 ggt gcg ggt gtg gat cca gca gac ttt gct gcc gat cca caa aag tac 2919
786 Gly Ala Gly Val Asp Pro Ala Asp Phe Ala Ala Asp Pro Gln Lys Tyr
789      895      900      905
795 gac atc cca gac tct gtc atc gcg ttc ctg cgc ggc gag ctt ggt aac 2967
798 Asp Ile Pro Asp Ser Val Ile Ala Phe Leu Arg Gly Glu Leu Gly Asn
801      910      915      920
807 cct cca ggt ggc tgg cca gag cca ctg cgc acc cgc gca ctg gaa ggc 3015
810 Pro Pro Gly Gly Trp Pro Glu Pro Leu Arg Thr Arg Ala Leu Glu Gly
813      925      930      935
819 cgc tcc gaa ggc aag gca cct ctg acg gaa gtt cct gag gaa gag cag 3063
822 Arg Ser Glu Gly Lys Ala Pro Leu Thr Glu Val Pro Glu Glu Glu Gln
825 940      945      950      955
831 gcg cac ctc gac gct gat tcc aag gaa cgt cgc aat agc ctc aac 3111
834 Ala His Leu Asp Ala Asp Asp Ser Lys Glu Arg Arg Asn Ser Leu Asn
837      960      965      970
843 cgc ctg ctg ttc ccg aag cca acc gaa gag ttc ctc gag cac cgt cgc 3159
846 Arg Leu Leu Phe Pro Lys Pro Thr Glu Glu Phe Leu Glu His Arg Arg
849      975      980      985
855 cgc ttc ggc aac acc tct gcg ctg gat gat cgt gaa ttc ttc tac ggc 3207
858 Arg Phe Gly Asn Thr Ser Ala Leu Asp Asp Arg Glu Phe Phe Tyr Gly
861      990      995      1000
867 ctg gtc gaa ggc cgc gag act ttg atc cgc ctg cca gat gtg cgc acc 3255
870 Leu Val Glu Gly Arg Glu Thr Leu Ile Arg Leu Pro Asp Val Arg Thr
873      1005      1010      1015
879 cca ctg ctt gtt cgc ctg gat gcg atc tct gag cca gac gat aag ggt 3303
882 Pro Leu Leu Val Arg Leu Asp Ala Ile Ser Glu Pro Asp Asp Lys Gly
885 1020      1025      1030      1035
891 atg cgc aat gtt gtg gcc aac gtc aac ggc cag atc cgc cca atg cgt 3351
894 Met Arg Asn Val Val Ala Asn Val Asn Gly Gln Ile Arg Pro Met Arg
897      1040      1045      1050
903 gtg cgt gac cgc tcc gtt gag tct gtc acc gca acc gca gaa aag gca 3399
906 Val Arg Asp Arg Ser Val Glu Ser Val Thr Ala Thr Ala Glu Lys Ala
909      1055      1060      1065
915 gat tcc tcc aac aag ggc cat gtt gct gca cca ttc gct ggt gtt gtc 3447
918 Asp Ser Ser Asn Lys Gly His Val Ala Ala Pro Phe Ala Gly Val Val
921      1070      1075      1080
927 acc gtg act gtt gct gaa ggt gat gag gtc aag gct gga gat gca gtc 3495
930 Thr Val Thr Val Ala Glu Gly Asp Glu Val Lys Ala Gly Asp Ala Val
933      1085      1090      1095
939 gca atc atc gag gct atg aag atg gaa gca aca atc act gct tct gtt 3543
942 Ala Ile Ile Glu Ala Met Lys Met Glu Ala Thr Ile Thr Ala Ser Val
945 1100      1105      1110      1115
951 gac ggc aaa atc gat cgc gtt gtg gtt cct gct gca acg aag gtg gaa 3591
954 Asp Gly Lys Ile Asp Arg Val Val Val Pro Ala Ala Thr Lys Val Glu
957      1120      1125      1130
963 ggt ggc gac ttg atc gtc gtc gtt tcc taa 3621
966 Gly Gly Asp Leu Ile Val Val Val Ser

```

Use of n and/or Xaa has been detected in the Sequence Listing.
 Review the Sequence Listing to insure a corresponding
 explanation is presented in the <220> to <223> fields of
 each sequence using n or Xaa.

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/045,072

DATE: 01/28/2002

TIME: 14:08:18

Input Set : A:\seqlist-15330790002 ascii

Output Set: N:\CRF3\01282002\J045072.raw

L:33 M:270 C: Current Application Number differs, Replaced Current Application No

L:33 M:271 C: Current Filing Date differs, Replaced Current Filing Date

L:2784 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:38